The Plant Journal (2008) **53**, 854–863 doi: 10.1111/j.1365-313X.2007.03366.x

The pyruvate, orthophosphate dikinase regulatory proteins of Arabidopsis possess a novel, unprecedented Ser/Thr protein kinase primary structure

Chris J. Chastain[1,†,*], Wenxin Xu[2,†], Kate Parsley[3], Gautam Sarath[4], Julian M. Hibberd[3] and Raymond Chollet[2]
[1]*Department of Biosciences, Minnesota State University-Moorhead, Moorhead, MN 56563, USA,*
[2]*Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, NE 68588-0664, USA,*
[3]*Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK, and*
[4]*USDA-ARS Grain, Forage & Bioenergy Research Unit, and University of Nebraska-Lincoln, Lincoln, NE 68583, USA*

Received 16 August 2007; revised 1 October 2007; accepted 19 October 2007.
[*]For correspondence (fax 1 218 477 2018; e-mail chastain@mnstate.edu).
[†]These authors contributed equally to this work.

Summary

Pyruvate, orthophosphate dikinase (PPDK) is a ubiquitous, low-abundance metabolic enzyme of undetermined function in C3 plants. Its activity in C3 chloroplasts is light-regulated via reversible phosphorylation of an active-site Thr residue by the PPDK regulatory protein (RP), a most unusual bifunctional protein kinase (PK)/protein phosphatase (PP). In this paper we document the molecular cloning and functional analysis of the two unique C3 RPs in *Arabidopsis thaliana*. The first of these, *AtRP1*, encodes a typical chloroplast-targeted, bifunctional C4-like RP. The second *RP* gene, *AtRP2*, encodes a monofunctional polypeptide that possesses *in vitro* RP-like PK activity but lacks *PP* activity, and is localized in the cytosol. Notably, the deduced primary structures of these two highly homologous polypeptides are devoid of any canonical subdomain structure that unifies all known eukaryotic and prokaryotic Ser/Thr PKs into one of three superfamilies, despite the direct demonstration that AtRP1 is functionally a member of this group. Instead, these C3 RPs and the related C4 plant homologues encode a conserved, centrally positioned, approximately 260-residue sequence currently described as the 'domain of unknown function 299' (DUF 299). We propose that vascular plant RPs form a unique protein kinase family now designated as the DUF 299 gene family.

Keywords: Ser/Thr protein kinase, protein phosphatase, PPDK regulatory protein (PPDK-RP), pyruvate, orthophosphate dikinase (PPDK), *Arabidopsis thaliana*, C3 plant.

Introduction

Owing to its importance as a pivotal regulator of the C4 photosynthetic cycle, the recent initial cloning and preliminary characterization of the maize (*Zea mays*) C4 pyruvate, orthophosphate dikinase (PPDK) regulatory protein (RP) represented a significant breakthrough in the field of C4 photosynthesis research (Burnell and Chastain, 2006). The pivotal role of RP is to post-translationally adjust the level of PPDK activity, a rate-limiting C4 cycle enzyme (Chastain and Chollet, 2003; Furbank *et al.*, 1997), to the available light energy incident on the C4 leaf. The regulatory protein does this by catalysing the rapid, reversible dark-induced phosphorylation/inactivation of a specific PPDK active-site Thr residue (Thr456 in maize; Burnell and Hatch, 1983, 1985; Chastain and Chollet, 2003) (Figure 1). In addition to its well-known role in C4 plants, RP is also present in C3 leaves where it confers a similar light/dark regulation of chloroplast-localized PPDK (Chastain and Chollet, 2003; Chastain *et al.*, 2002), but unlike C4 dikinase, C3 PPDK plays no direct role in photosynthesis. C3 RP has also been demonstrated to occur in developing rice (*Oryza sativa*) seeds where it functions to irreversibly phosphorylate/inactivate PPDK during seed maturation (Chastain *et al.*, 2006). Beyond documenting its presence in C3 plants, the functional role(s) for PPDK, and its regulation by RP, have yet to be established in C3 species (Chastain and Chollet, 2003; Chastain *et al.*, 2006).

Long before the recent molecular cloning of a C4 RP, the collective biochemical properties of the C4 maize RP were discovered to be singularly unique for a regulatory enzyme

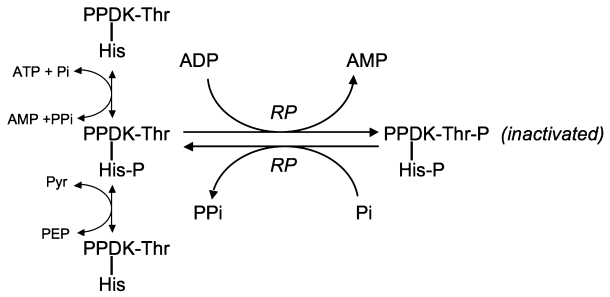

Figure 1. Pyruvate,orthophosphate dikinase (PPDK) catalysis and its phospho-regulation by regulatory protein (RP).
During formation of the essential His-P reaction intermediate of PPDK during catalysis (His-458 in maize; vertical), the enzyme is amenable to reversible inactivation by RP via ADP-dependent phosphorylation of a proximal Thr residue (Thr-456 in maize; horizontal). Regulatory protein also catalyses reactivation of PPDK-ThrP via a Pi-dependent, PPi-forming phosphorolytic reaction.

(Burnell and Hatch, 1983, 1985; Roeske and Chollet, 1987; and for a recent review see Chastain and Chollet, 2003). These collective novel properties include: (i) its bifunctionality, catalysing both PPDK phosphorylation/inactivation and dephosphorylation/reactivation; (ii) the use of ADP versus ATP as its phosphoryl donor; and (iii) its utilization of an inorganic phosphate (Pi) dependent, inorganic pyrophosphate (PPi) forming phosphorolytic dephosphorylation mechanism rather than the simple anhydride bond hydrolysis utilized by most protein phosphatases (PPs) (Figure 1). Additional insight into its enzyme properties were later gained by studies that selectively substituted Ser and Tyr for RP's target Thr residue in the maize PPDK active site (Chastain *et al.*, 1997, 2000). These site-specific substitutions revealed that Ser but not Tyr was functionally interchangeable with the wild-type (WT) Thr residue. The insightful implication of this observation was that RP was functionally and, by extrapolation, structurally related to the Ser/Thr superfamily of eukaryotic protein kinases (PKs) (Hanks and Hunter, 1995; Hardie, 1999; Scheeff and Bourne, 2005). In view of this, it was anticipated that candidate gene(s) for RP could be identified by bioinformatics and ultimately cloned based on its familial Ser/Thr PK domain structure. The premise for this is that all known eukaryotic Ser/Thr PKs share a canonical 12-subdomain primary structure, all of which are requisite for enzymatic phosphorylation of target Thr/Ser substrate residues (Hanks and Hunter, 1995; Hardie, 1999; Scheeff and Bourne, 2005). Even the most minimal Ser/Thr PK, the approximately 31-kDa PEP-carboxylase kinase (PpcK) of plants, encodes these 12 signature subdomains (Hartwell *et al.*, 1999; Xu *et al.*, 2003).

In this paper we detail the molecular cloning and complementary functional and bioinformatics analysis of the two C3 *RP* genes encoded by the Arabidopsis genome. Unexpectedly, we show herein that both deduced RPs, like their C4 maize RP counterpart (Burnell and Chastain, 2006), are devoid of any of the known canonical PK subdomains required for enzymatic function as a Ser/Thr kinase. We therefore propose that the C3 RPs and related C4 homologues represent a new structural paradigm for enzymes that catalyse regulatory phosphorylation of protein-hydroxyamino acids. Additionally, we document how these two contrasting RP gene products play distinctly different roles in regulating PPDK activity in C3 plants.

Results and discussion

Cloning and sequence determination of AtRP1/2

Because of the differing functions of PPDK in C4 versus C3 plants, we surmised that the respective C3 plant RP would also display regulatory functions divergent from C4 RP in those C3 plant organs where PPDK is expressed. To fully explore this hypothesis, we sought to clone a functionally representative C3 RP from the model C3 dicot Arabidopsis. To accomplish this, we initially settled on an interaction-based cloning approach as past attempts using more traditional methods had proved fruitless (Burnell and Chastain, 2006; Smith *et al.*, 1994a). During the development of the requisite cloning tools required for this strategy, a proteomics study was published that profiled differential expression of soluble stromal polypeptides in isolated maize leaf mesophyll and bundle sheath cell chloroplasts (Majeran *et al.*, 2005). In this study we noted, as did the authors, the presence of a low-abundance polypeptide of unknown function, specific to C4 mesophyll cell chloroplasts. The cross-listed Arabidopsis homologue to this maize polypeptide was At4g21210. Our bioinformatics analysis of its predicted polypeptide failed to reveal any canonical Ser/Thr PK subdomain structure other than a putative, centrally positioned eight-residue P-loop (Hulo *et al.*, 2006). Nevertheless, we further pursued this lone candidate gene since the well established genomic Arabidopsis resources permitted us to rapidly isolate its full-length cDNA for subsequent facile testing of the encoded polypeptide for *in vitro* RP activity. In parallel with our efforts in cloning the C3 RP gene, a collaborative effort with Dr J. N. Burnell and co-workers (James Cook University, Queensland, Australia), succeeded in the seminal cloning of the C4 RP gene from maize (Burnell and Chastain, 2006), which was the same maize polypeptide profiled in the proteomics study that was cross-listed with At4g21210 (Majeran *et al.*, 2005). This discovery led to our confirming, by similar means, that *At4g21210* encoded a bona fide bifunctional C3 RP, designated AtRP1 (see the following section). We then utilized the deduced AtRP1 amino acid (aa) sequence for searching the Arabidopsis nuclear genome for homologous genes and identified a second RP-like gene, *At3g01200*. Largely due to this chronological order of identification, we termed At3g01200 as AtRP2. The aligned AtRP1/2 primary structures as deduced

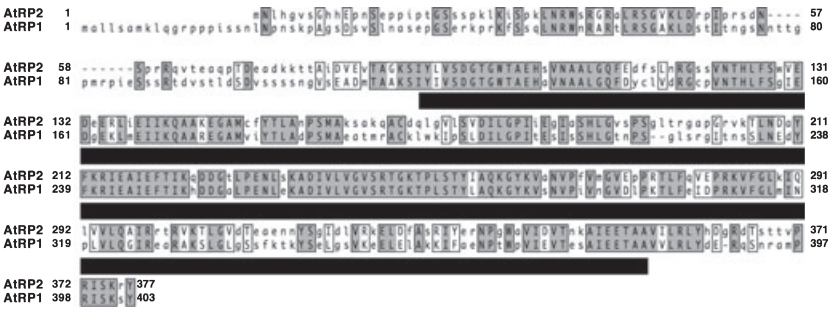

Figure 2. Direct alignment of the full-length, deduced amino-acid (aa) sequences of AtRP1/2. Residues identified as corresponding to the conserved DUF 299 domain (Hulo *et al.*, 2006) are underscored (bar below). Overall, AtRP1/2 share a 67% aa sequence identity (exact, 54%; similar, 13%).

from the respective isolated full-length cDNA clones (Figure 2) were found to be in exact agreement with the data base sequence in predicting a 403-aa, 43.7 kDa pre-protein for AtRP1 and a 377-residue, 41.4-kDa polypeptide for AtRP2. A dominating structural element of the Arabidopsis RP polypeptides, as in maize C4 RP (Burnell and Chastain, 2006), is a centrally positioned domain of unknown function (DUF 299; Hulo *et al.*, 2006) (Figure 2). Direct alignment of the two AtRP polypeptides shows a strong sequence divergence at the N-terminal portion of the polypeptide (Figure 2). However, as only *AtRP1* is predicted to encode a chloroplast transit peptide (Emanuelsson *et al.*, 1999; Small *et al.*, 2004; and see data below), just how divergent the two mature polypeptides are at their respective N-termini is at this time uncertain.

AtRP1 is a bifunctional, C4-type RP while AtRP2 is a monofunctional, PK-only RP variant

Two complementary RP assay methods were used to assess the *in vitro* activity of recombinant AtRP1/2. The first procedure consisted of an immuno-based assay of site-specific phosphorylation and dephosphorylation of the regulatory Thr residue of Arabidopsis PPDK (Chastain *et al.*, 2000, 2002). A second method utilized a spectrophotometric-based assay of PPDK enzyme activity before and after incubation with exogenous RP (Roeske and Chollet, 1987). In the immuno-based assay, both AtRP1 and AtRP2 were shown to catalyse ADP-dependent, site-specific threonyl phosphorylation of dephospho-PPDK (Figure 3a). As noted in Figure 3a, ATP was included along with ADP in the RP PK reaction since the requisite RP phosphorylation substrate is the transient PPDK-His-P catalytic reaction intermediate (Figure 1). This is substantiated by the finding that phosphorylation of PPDK by AtRP1/2 is negated or greatly reduced when pyruvate is also included in the reaction mixture (Figure 3a). The added pyruvate in effect scavenges the catalytic His-P reaction intermediate during conversion to PEP (Figure 1), thus removing the PPDK-His-P enzyme RP substrate. Thus, both recombinant AtRP1/2 display the unique RP PK substrate specificity (i.e. ADP-dependent, PPDK-His-P-dependent). When the Pi-dependent PP activity

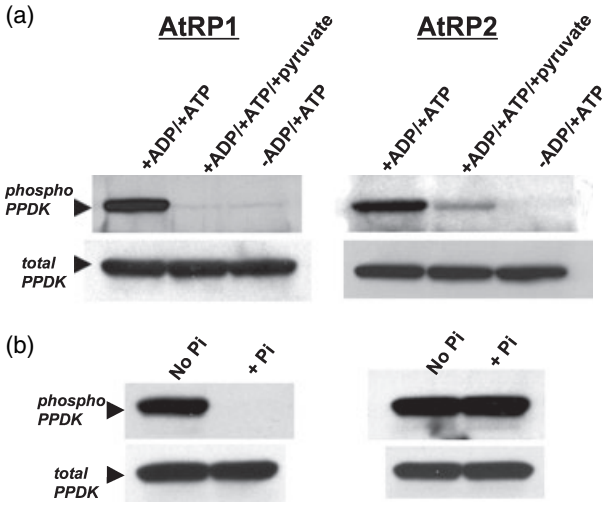

Figure 3. Immuno-based assays of (a) Arabidopsis regulatory protein (AtRP) protein kinase (PK) and (b) AtRP protein phosphatase (PP) activity.
Shown are representative denaturing Western blots of quenched, 10-μl assay reaction aliquots probed with anti-pyruvate,orthophosphate dikinase (PPDK)-ThrP or anti-PPDK polyclonal antibodies as described in Experimental procedures. Noted above each lane are variations in the standard reaction mixture: +ADP (1 mM); +ATP (0.2 mM); +pyruvate (2 mM); +Pi (2.5 mM). Arrowheads indicate position of bands corresponding to the approximately 94-kDa Arabidopsis PPDK monomer as estimated by molecular mass standards on the same blot.

of RP was assayed, AtRP1 was shown to reduce the initial amount of PPDK-ThrP at time 0 to immunologically undetectable levels after 20 min (Figure 3b). Unexpectedly, AtRP2 did not catalyse a Pi-dependent decrease in the initial amount of phospho-PPDK over the same 20-min incubation period (Figure 3b).

We then proceeded to corroborate the results of the immuno-based RP assays with the corresponding spectrophotometric-based, RP-catalysed PPDK inactivation/reactivation assay. Overall, the results from these assays were in excellent agreement with the findings of the immuno-based RP assays. For the RP PK reaction, both recombinant AtRPs were shown to catalyse the ADP-dependent inactivation of active, dephospho-PPDK (Table 1a). Inclusion of pyruvate in the reaction, as in the immuno-based assays (Figure 3a), greatly negated this inactivation/phosphorylation. However,

a notable difference between AtRP1 and -2 was in the extent of the inactivation of dephospho-PPDK, with AtRP1 catalysing a nearly complete (99%) inactivation of initial PPDK activity versus a 43% inactivation by AtRP2. Since these values represent only an initial analysis of the functional RP properties, we hesitate at this time to speculate on the possible biochemical significance of this observation. In the corresponding RP PP assays (Table 1a), AtRP1 demonstrated a strict, Pi-dependent reactivation of inactive phospho-PPDK (0.9% to 70%). Conversely, AtRP2 did not catalyse reactivation of inactive phospho-PPDK, thus corroborating the RP immuno-based assay results (Figure 3b).

Given the strict conservation of the active-site residues surrounding the regulatory Thr and proximal, essential His among vascular plant PPDKs, including maize and Arabidopsis (Chastain and Chollet, 2003), and the documented cross-functionality of RP for heterologous PPDKs (Chastain and Chollet, 2003; Chastain *et al.*, 2002; Smith *et al.*, 1994b), we utilized maize C4 PPDK active-site Thr-456 mutants (Chastain *et al.*, 1997) as substrates for testing AtRP1 PK Thr–Ser equivalency. In the spectrophotometric-based inactivation assay, AtRP1 was shown to be highly effective in catalysing inactivation (approximately 99%) of active/dephospho WT maize PPDK (Table 1b). Notably, AtRP1 catalysed a comparable ADP-dependent inactivation of maize PPDK when Ser was substituted for the WT Thr456 target residue (T456S mutant, Table 1b). With a non-phosphorylatable Val at this same position (T456V) AtRP1 was unable to catalyse any inactivation of active C4 PPDK. Lastly, we investigated the ability of AtRP1 to reactivate inactivated/phosphorylated maize WT and T456S PPDK. Interestingly, these assays showed that while AtRP1 could catalyse a robust, Pi-dependent reactivation of inactive WT C4 PPDK, it was incapable of catalysing reactivation of inactive T456S PPDK (Table 1b). Presumably, unlike the RP PK function, which can utilise either Thr or Ser as a phosphorylatable target (Chastain *et al.*, 1997), the RP PP function has a strict substrate requirement for the threonyl phosphate and is unable to catalytically dephosphorylate Ser-P at this same position.

Table 1. Spectrophotometric-based RP PK- and PP-assays. Assay of recombinant (a) AtRP1 and AtRP2 using *Arabidopsis* PPDK as substrate, and (b) AtRP1 using maize WT and Thr-456 mutant recombinant PPDKs as substrate. Values displaying a 43% or greater change were determined to be statistically significant ($P < 0.001$) using the two-sample *t* test (95% confidence level). Number of independent determinations (*n*) for the above data are: (a) AtRP1 PK assays: +ADP, $n = 7$; −ADP, $n = 8$; +ADP/+pyruvate, $n = 5$; AtRP2 PK assays: +ADP, $n = 8$; −ADP, $n = 4$; +ADP/+pyruvate $n = 5$; AtRP1 PP assays: +Pi, $n = 3$; −Pi, $n = 3$; AtRP2 PP assays: +Pi, $n = 3$; −Pi, $n = 3$; (b) $n = 3$ for all data.

(a)	Arabidopsis PPDK activity (μmol min^{-1}mg^{-1}protein^{-1})	
PK assay (additions)	Initial activity	+20 min incubation
+AtRP1/−ADP	4.10	4.80
+AtRP1/+ADP	4.38	0.03
+AtRP1/+ADP/+Pyruvate	3.42	3.03
+AtRP2/−ADP	4.44	4.16
+AtRP2/+ADP	4.17	2.39
+AtRP2/+ADP/+Pyruvate	3.94	3.60
PP assay (additions)		
+AtRP1/−Pi	0.05	0.03
+AtRP1/+Pi	0.04	2.89
+AtRP2/−Pi	∼0	0.01
+AtRP2/+Pi	0.02	∼0

(b)		Maize PPDK activity (μmol min^{-1}mg prot^{-1})	
Maize PPDK-type (as AtRP1 substrate)	PK assay (additions)	Initial activity	+20 min incubation
WT	−ADP	4.40	4.28
WT	+ADP	4.42	0.03
T456S	−ADP	3.71	3.64
T456S	+ADP	3.57	0.05
T456V	−ADP	4.26	4.54
T456V	+ADP	4.40	4.33
WT	−Pi	0.04	0.05
WT	+Pi	0.05	4.50
T456S	−Pi	0.08	0.07
T456S	+Pi	0.02	0.06

AtRP1 is chloroplast-targeted while AtRP2 is cytosol-localized

To gain initial insight into the intracellular localization of AtRP1/2, the online ChloroP (Emanuelsson *et al.*, 1999) and Predotar (Small *et al.*, 2004) prediction algorithms were used to interrogate the deduced aa sequences for putative transit-peptide encoding sequences. ChloroP predicted that both AtRP polypeptides contained N-terminal chloroplast-targeting transit peptides of 86 and 69 residues for AtRP1/2, respectively. Alternatively, the Predotar analysis predicted AtRP1 to be 'possibly' plastid targeted, and the AtRP2 polypeptide to lack a targeting transit peptide. To experimentally resolve these rather ambiguous predictions, we produced a series of AtRP1/2 C-terminal green fluorescent protein (GFP) translational fusions plus an additional AtRP2 N-terminal cyan fluorescent protein (CFP) translational fusion to observe *in planta* the subcellular localization of each respective AtRP using confocal laser scanning microscopy. After microprojectile bombardment of full-length AtRP–GFP fusions constructs into C3 tobacco leaves, we observed that the *AtRP1* full-length open reading frame (ORF)-directed GFP accumulation into chloroplasts (Figure 4a–d), while the *AtRP2* full-length ORF generated cytosolic GFP accumulation (Figure 4i–l). Proteomic analysis has indicated that many proteins lacking *in silico*-predicted N-terminal transit peptides may also accumulate in plastids (Friso *et al.*, 2004; Peltier *et al.*, 2002). To rule out the possibility that a GFP fusion at the

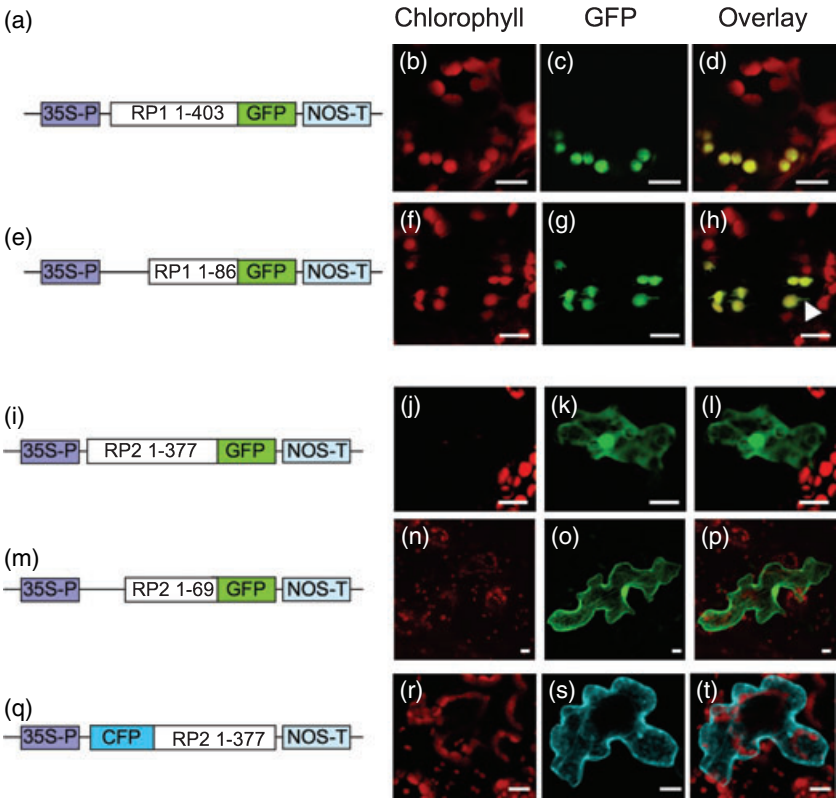

Figure 4. Microprojectile bombardment followed by confocal laser scanning microscopy indicates Arabidopsis regulatory protein 1 (AtRP1) is chloroplast-targeted while AtRP2 is cytosol-localized.
Image sets b–d, f–h and j–l are from tobacco leaves, while image sets n–p and r–t are from Arabidopsis leaves. Adjacent construct diagram key: 35S-P, the CaMV-35S promoter; RP1 and RP2 indicate the respective *AtRP* open reading frame; CFP, cyan fluorescent protein; NOS-T, nopaline synthase terminator.
(b, f, j, n, r) Images from the chlorophyll channel, where red represents chlorophyll autofluorescence.
(c, g, k, o, s) Images from the GFP (CFP) channel, where green (aqua-blue) represents GFP (CFP).
(d, h, l, p, t) Merged images from the chlorophyll and GFP (CFP) channels. Scale bars represent 10 μm, and the arrowhead in panel (h) indicates a chloroplast stromule.

C-terminus of AtRP2 led to aberrant accumulation of GFP, we generated a corresponding N-terminal CFP–AtRP2 fusion construct for subsequent transformation into Arabidopsis leaf cells. As shown in Figure 4q–t, this fusion construct also directed cytosolic accumulation of CFP, providing additional evidence that AtRP2 is native to this subcellular compartment. We then tested GFP constructs for which only the extreme, N-terminal AtRP1/2 sequences, predicted by ChloroP as encoding chloroplast transit peptides, were translationally fused to GFP. These results, shown in Fig. 4e–h and 4m–p, respectively, revealed that the predicted N-terminal 86 aa of AtRP1 were sufficient for directing accumulation of the GFP fusion protein in C3 chloroplasts (tobacco), while the corresponding AtRP2 construct (N-terminal 69 aa residues) led to GFP accumulation in the cytosol (Arabidopsis). Overall, these results indicate that AtRP1, like its C4 maize RP homologue (Smith *et al.*, 1994a), is targeted to the chloroplast stroma, while AtRP2 is cytosolic.

The C4-like AtRP1 is the predominant RP transcript in green tissues, while the PK-only variant AtRP2 transcript is prevalent in pollen and seeds

In order to begin to assess putative functions for the two divergent AtRPs we profiled organ-specific *AtRP1/2* and *AtPPDK* transcript levels using the Affymetrix 22k ATH1 Arabidopsis genome array chip via the online Genevestigator-Gene Atlas (Zimmermann *et al.*, 2004). These data revealed that the *AtRP1* transcript is predominant in greening and green tissues and organs (Figure 5). This is consistent with the finding that AtRP1 is chloroplast-targeted (Figure 4). Conversely, the *AtRP2* transcript is relatively low in green tissues (approximately 20% of *AtRP1*), moderately higher in seeds, and shows a striking peak level in pollen, where the corresponding transcript for *PPDK* is also maximal (Figure 5). In contrast, *AtRP1* transcript in pollen appeared to be essentially repressed. It should be noted, however, that due to the unusual dual-promoter structure of the single Arabidopsis *PPDK* gene (Parsley and Hibberd, 2006), the microarray analysis does not differentiate between plastid-targeted and cytosol-targeted *PPDK* transcripts. Therefore, it is not known at that time if PPDK in pollen is largely cytosolic and co-localized with AtRP2, but this seems likely given the low plastid content of this organ. Transcript levels in other plant organs from the same array data set, such as roots, stems, and vascular tissue, showed only minimal mRNA levels for these three genes and thus were not selected for inclusion in Figure 5 (Zimmermann *et al.*, 2004).

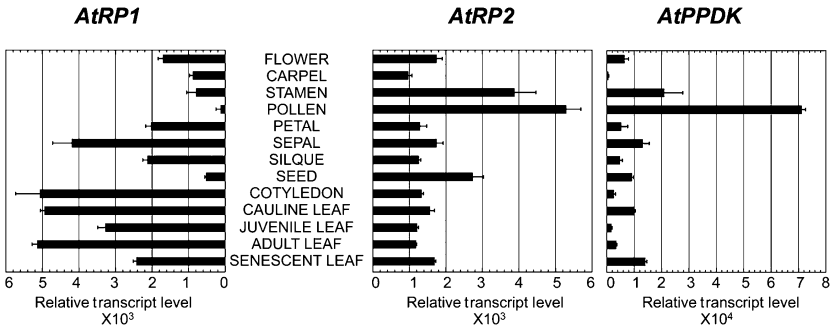

Figure 5. Relative transcript abundance of *AtRP1* (*At4g21210*), *AtRP2* (*At3g01200*) and *AtPPDK* (*At4g15530*) in select wild-type (WT) Arabidopsis organs.
Data were obtained by performing on-line Genevestigator-Gene Atlas queries of the Affymetrix 22k ATH1 Arabidopsis genome array chip (http://https://www.genevestigator.ethz.ch/at/; Zimmermann *et al.*, 2004). Note the 10-fold higher unit *x*-axis used for the *AtPPDK* transcript panel.

Bioinformatics analysis reveals AtRP1/2 and the allied DUF 299 family members are devoid of canonical PK subdomain structure and lack defined PP motifs

As discussed above, we and others (Burnell and Chastain, 2006) perceived that the primary aa sequence of plant RP would reveal a canonical PK subdomain structure as related to its eukaryotic Ser/Thr PK activity as well as motif structure related to its unusual PP function. We initially searched for such elements by performing refined, trial-by-trial bioinformatics interrogations of deduced AtRP1/2 polypeptide sequence using all the publicly available algorithms but failed to reveal even weakly facsimile eukaryotic or prokaryotic PK subdomain structure (e.g. PROSITE, Swiss-Prot, MotifScan). A more rigorous analysis was then executed by performing an algorithm-aided custom alignment of AtRP1/2 with representative (AtCDPK4, AtCDPK13) and atypical (AtPpcK1, AtPpcK2) plant Ser/Thr PKs that would allow us to scrutinize these RP polypeptides for putative PK subdomain substructure and/or key residues with exacting precision (Figure 6a). In this explicit alignment, we could not locate any primary structure within the

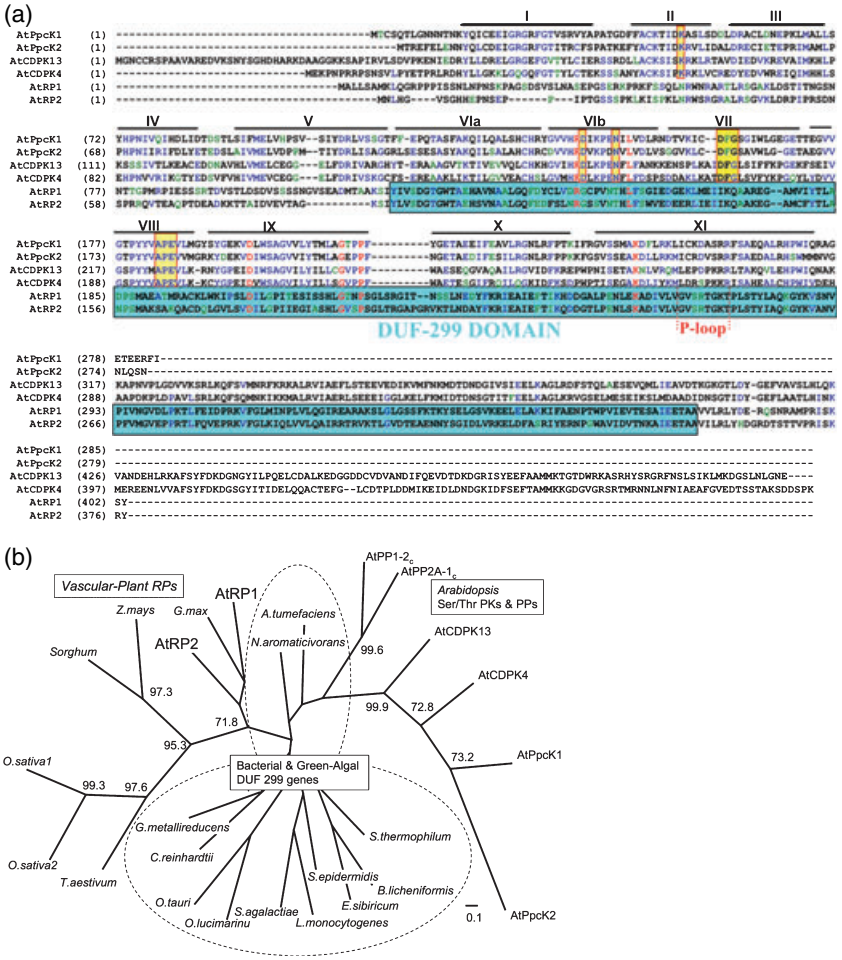

Figure 6. Multiple alignment and molecular phylogenetic analysis of AtRP1/2 deduced amino-acid sequences.
(a) Multiple alignment of the Arabidopsis regulatory protein AtRP1/2 deduced sequences with typical plant Ser/Thr protein kinases (PKs) (AtCDPK4, AtCDPK13), and atypical plant Ser/Thr PKs (AtPpcK1, AtPpcK2; Hartwell *et al.*, 1999; Xu *et al.*, 2003, 2007). The 12 canonical, conserved catalytic Ser/Thr-kinase subdomains I–XI (Hanks and Hunter, 1995; Hardie, 1999; Scheeff and Bourne, 2005) are shown positioned above the corresponding portions of the aligned PK sequences. The invariant and catalytically essential Ser/Thr kinase residues are boxed in yellow. The DUF 299 domain and its centrally positioned, putative P-loop in AtRP1/2 are boxed in aqua-blue and red, respectively.
(b) Phylogenetic tree analysis of AtRP1/2 with: (i) vascular plant and green algal RP homologues (all putative RPs except for *Zea mays*; Burnell and Chastain, 2006); (ii) bacterial DUF 299/RP-like genes; (iii) Arabidopsis Ser/Thr PKs (from panel a); and (iv) Arabidopsis PP1-2_c and PP2A-1_c catalytic subunits. Full-length open reading frames for the various species used in the tree analysis were obtained from GenBank or from our unpublished data (Table S1). The phylogenetic tree was constructed with the PHYLIP software package based on the alignments using the CLUSTALX program, and visualized by TREEVIEW 32 software. The numbers at the nodes are bootstrap supporting values (%) for 1000 replications. Only values >70% are shown.

AtRP1/2 polypeptides or the internal approximately 260-residue DUF 299 domain that would correlate with the canonical subdomains I–XI inherent in all known eukaryotic Ser/Thr PKs (Hanks and Hunter, 1995; Hardie, 1999; Scheeff and Bourne, 2005), or the catalytically essential (and invariant) Ser/Thr PK residues (Figure 6a, yellow-boxed). We then performed an unrooted molecular phylogenetic analysis with the full-length AtRP1/2 aa sequences against select bacterial DUF 299 genes, vascular plant RPs and representative Arabidopsis Ser/Thr PKs and Ser/Thr PPs. This analysis, shown in Figure 6b, supports our conclusions from the related alignment analysis that RP, whether from plants, green algae or the putative bacterial RPs (as DUF 299-containing genes), is unrelated to the canonically structured plant Ser/Thr PKs. This tree analysis also demonstrates that the PP activity inherent in the AtRP1 primary structure (Table 1, Figure 3b) is apparently highly divergent from the ubiquitous PP1- and PP2A-catalytic subunits included in the analysis.

Concluding remarks

In this paper we have demonstrated that two functionally and spatially distinct RPs are present in C3 plant cells, rather than a single plastidic RP as was previously presumed (Chastain and Chollet, 2003; Chastain *et al.*, 2002). We document that the first of these, AtRP1, corresponds to the previously characterized C4-like, plastid-localized RP (Chastain and Chollet, 2003; Chastain *et al.*, 2002), having typical C4-like bifunctional PK/PP properties; the second RP, AtRP2, we propose to be a monofunctional, PK-only variant that is restricted to the cytosolic compartment. The occurrence of two functionally unique C3 RPs, each targeted at different subcellular pools of PPDK, has important consequences for how dikinase may participate in ancillary metabolic pathways. In C3 plants, the metabolic role(s) of PPDK has remained an enigma (Chastain and Chollet, 2003; Chastain *et al.*, 2006). One hypothesis concerning its function is that PPDK, like other ubiquitous 'housekeeping' metabolic enzymes, is deployed in yet to be elucidated 'compensatory' metabolic pathways related to plant biotic stress or development (Dennis and Blakeley, 2000; Plaxton and Podesta, 2006). Moreover, because catalysis by PPDK is readily reversible (Chastain and Chollet, 2003), it may be a versatile enzyme for participating in one of a number of such metabolic 'backup' or 'bypass' reactions (Chastain and Chollet, 2003; Chastain *et al.*, 2006; Lasanthi-Kudahettige *et al.*, 2007; Méchin *et al.*, 2007; Plaxton and Podesta, 2006). Prior to the present work it was presumed that cytosolic PPDK was not subject to regulatory phosphorylation and thus rendered constitutively active, thereby positing its putative role in a particular metabolic process as problematic (Lasanthi-Kudahettige *et al.*, 2007; Méchin *et al.*, 2007). However, with the knowledge that a C3 RP variant is co-localized with C3 PPDK in the cytosolic compartment, the latter's coordinate regulation with opposing metabolic processes may be more easily rationalized. Further clues in this regard are provided by the microarray data that show a striking correlation of peak transcript abundance for *PPDK* and *AtRP2* in pollen, leading us to conjecture that AtRP2 may play an important role in pollen development and/or function by regulatory inactivation of PPDK.

Of most profound significance is the finding that the deduced primary structures encoded by the two Arabidopsis *RP* genes represent a hitherto unknown family of eukaryotic or prokaryotic Ser/Thr kinases. Our bioinformatics analyses of the RP polypeptides show that these proteins are devoid of any catalytic subdomain structure that unifies all known eukaryotic and prokaryotic PKs into one of three superfamilies (Hanks and Hunter, 1995; Hardie, 1999; Leonard *et al.*, 1998; Scheeff and Bourne, 2005). Rather, we propose that the vascular plant RPs form a unique PK family, now designated as the DUF 299 gene family. By definition, the DUF designation is assigned to conserved aa-encoding sequences that are recurrent in various genomic databases, but have no known functional precedent. In the two AtRPs, this approximately 260-aa domain spans the central core of the 377-residue AtRP2 and 403-residue AtRP1 full-length polypeptides. As in the recently reported maize RP gene product (Burnell and Chastain, 2006), the only bioinformatics-deciphered motif structure for the AtRPs' DUF 299 is a centrally positioned, eight-residue P-loop (Figure 6a). Interestingly, this putative new family of protein-phosphorylating enzymes is restricted phylogenetically to vascular plants, green algae and a diverse group of PPDK-encoding prokaryotes (Whelan *et al.*, 2006) (Figure 6b). Curiously, the only non-plant eukaryote that has been identified by bioinformatics to contain a DUF 299 signature is *Anopheles gambia*, but this arthropod shows the weakest homology of all the described plant and bacterial DUF 299 genes, and may thus be artefactual. Overall, our molecular phylogenetic analysis of database-mined plant and bacterial RP sequences suggests a green algal origin for the vascular plant RPs. In addition, we were unable to identify by bioinformatics another plant species besides Arabidopsis that possesses more than one distinct *RP* gene, although the single-locus rice gene encodes an alternative translational start codon. However, given the apparent rarity of the *AtRP2* transcript in most Arabidopsis organs (Zimmermann *et al.*, 2004; Figure 5) and the paucity of fully annotated plant genomes, it is too early to draw conclusions on the species distribution of such a 'second-locus' *RP* gene among vascular plants. Lastly, we were not able to determine any distinct PP-related motif structure (Bhaduri and Sowdhamini, 2005; Stern *et al.*, 2007) in AtRP1 despite its documented bifunctionality, although it is possible that a more extensive sequence microanalysis may reveal that certain key residues are present (GS, personal communication). A far more chal-

lenging endeavour will be to elucidate the structural basis for the unusual AtRP1/2 and ZmRP1 (Burnell and Chastain, 2006) protein-phosphorylation mechanism and substrate requirements. Largely due to the absence of any reference structure for this novel PK, except for its putative P-loop, there is nothing upon which to base an empirical investigation into its reaction mechanism. Thus, understanding the catalytic mechanism of this novel, enigmatic family of Ser/Thr kinases will first require a lucid three-dimensional crystal structure before an empirical molecular dissection can yield meaningful information.

A final point to be made relates to the broad importance of regulatory protein kinases and protein phosphatases to all branches of life and biomedicine. That a new family of protein-phosphorylating enzymes has persisted as genes of unknown function until the present illustrates the potential for discovering yet more important regulatory PKs or PPs represented in the databases.

Experimental procedures

Isolation of full-length cDNAs for AtRP1/2

Full-length cDNAs were isolated from WT *Arabidopsis thaliana* (Columbia-0) rosette leaf mRNA using a RT-PCR procedure. Briefly, total RNA was isolated from 3- to 4-week-old leaves using an RNeasy Plant mini-kit (Qiagen, http://www1.qiagen.com/) for use as a template in synthesizing first-strand cDNA via an iScript cDNA synthesis kit (Bio-Rad, http://www.bio-rad.com/). A PfuUltra® High-Fidelity DNA polymerase kit (Stratagene, http://www.stratagene.com/) was then used to amplify the cDNA's corresponding ORF for *AtRP1* (1212 bp) or *AtRP2* (1134 bp) using the gene-specific primer pairs plus encoded restriction enzyme sites (as labelled): F-AtRP1-*Nde*I (5′-GGAATTCCATATGGCTTTGCTCTCGGCGATG-3′) + R-AtRP1-*Bam*HI (5′-GCGGATCCTTAGTAGCTTTTAGAGATGCGAGGC-3′); and F-AtRP2-*Nde*I (5′-GGATTAGCATATGAATTTACACGGCGTCTCGG-3′) + R-AtRP2-*Bam*HI (5′-GCGGATCCTTAGTAGCGTTTTGAGATAC-3′). The resulting PCR products were digested with *Nde*I and *Bam*HI and ligated into an *Nde*I, *Bam*HI linearized pET 28a plasmid (Novagen, http://www.emdbiosciences.com/), which fuses an N-terminal 6×His tag-10-aa enterokinase site 5′ to the native translational start codon of the respective *AtRP* ORF. Sequence fidelity of the AtRP-pET 28a constructs was authenticated against database sequences by DNA sequencing.

Affinity purification of recombinant AtRP1/2

The respective recombinant pET 28a AtRP1/2 protein expression plasmids were transformed into *Escherichia coli* BL21 DE3 host strains for subsequent use in large-scale culture and production of recombinant AtRP1/2 as follows. Ten millilitres of cells, harbouring an individual AtRP construct, was grown to saturation and used to inoculate 1 l flasks of Terrific Broth containing 100 μg ml^{-1} kanamycin. The cultures were then placed onto an incubator–shaker at 37°C and allowed to grow until the cell density reached an OD_{600} of approximately 0.6. At this time, transcription of recombinant RP from the pET 28a plasmids was induced by addition of 0.5 mM isopropyl-beta-D-thiogalactopyranoside (IPTG), and growth continued for another 3 h. The induced cells were harvested by centrifugation (5 min, 6000***g***) and the pellets washed once with 10 mM KPi, pH 8.0, 1 mM $MgCl_2$. Extraction of soluble protein was accomplished by resuspending cell pellets in 20 ml of binding buffer consisting of 50 mM KPi, pH 8.0, 5 mM imidazole, pH 8.0, 200 mM KCl, 2.5 mM $MgSO_4$, 5 mM 2-mercaptoethanol and one tablet of Complete®-minus-EDTA protease inhibitor cocktail (Roche Applied Science, http://www.roche-applied-science.com/). Cell lysis was accomplished by decanting resuspended cells into a 25-ml chambered bead-beater homogenizer (Bio-Spec Products, http://www.biospec.com/) plus 10 ml (dry volume) of 0.1 mm quartz beads. The resulting cell lysate was clarified of insoluble matter by high-speed centrifugation (20 min, 74 000***g***) and then combined with 1.5 ml Ni-NTA Superflow® Ni^{2+}-agarose beads (Qiagen) for batch-binding of 6×His-tagged recombinant proteins at 4°C for 2 h. The batch-bound slurry was subsequently decanted into a 1 cm diameter mini-column for sequentially washing the beads with 20 ml of wash buffer A + Tween-20 (50 mM KPi, pH 8.0, 30 mM imidazole, pH 8.0, 200 mM KCl, 2.5 mM $MgSO_4$, 5 mM 2-mercaptoethanol, 1% Tween-20, one tablet of Complete-minus-EDTA protease inhibitor cocktail), and then 20 ml of wash buffer A alone (minus 1% Tween-20). Recombinant 6×His-bound protein was eluted from the beads with wash buffer A except that the imidazole concentration was increased from 30 mM to 150 mM. The protein eluate was brought to 70% ammonium sulphate saturation and placed on wet ice overnight to maximize protein precipitation. After adding 2-mercaptoethanol (final, 20 mM), the overnight precipitate was collected by centrifugation (20 min, 32 000***g***) and subsequently resuspended in 2.5 ml of binding buffer, desalted via gel filtration and combined with 1.5 ml Ni-NTA Superflow agarose beads for batch-binding as above. The slurry was then poured into a 1 cm diameter Bio-Rad Econo column, washed with 10 ml of binding buffer, and the column attached in-line to an automated chromatography system (Bio-Rad Biologic System). The column was further washed with binding buffer until the A_{280} decreased to approximately 0. Recombinant AtRP protein was then eluted from the column using a linear, 50-ml 10–200 mM imidazole gradient in 50 mM KPi, pH 8.0, 200 mM KCl, 2.5 mM $MgSO_4$, 5 mM 2-mercaptoethanol and 1 mM phenylmethylsulphonyl fluoride (PMSF). Fractions (1 ml) were collected and assessed for AtRP purity via Coomassie stained SDS-PAGE gels. Fractions free of contaminating polypeptides (Figure S1) were pooled, brought to 70% ammonium sulphate saturation, and allowed to precipitate on wet ice overnight as above. Following the addition of 2-mercaptoethanol to 20 mM, 1-ml aliquots of this precipitate were transferred into 1.5-ml microfuge tubes, and the precipitated protein pelleted via centrifugation at 14 000***g***. Drained pellets were then flushed with N_2, sealed, and stored at −80°C. For use in biochemical assays, the stored affinity-purified AtRP was resuspended in a desalting buffer consisting of 50 mM Bicine, pH 8.3, 5 mM $MgCl_2$, 0.1 mM EDTA, 5 mM DTT and 1 mg ml^{-1} blue dextran and rapidly desalted using a Zeba® spin-column (Pierce Biotechnology, http://www.piercenet.com/).

In vitro *RP PK and PP assays*

Regulatory protein PK and PP activities were assayed using a sequential two-step procedure as modified from a previously described method (Chastain *et al.*, 2002; Roeske and Chollet, 1987). The first-step RP PK assay was executed by addition of affinity-purified AtRP (Figure S1) to a final concentration of 0.45 μg ml^{-1} into a reaction mixture consisting of 70 μg of affinity-purified recombinant Arabidopsis or maize PPDK (fully activated, non-phospho form), 50 mM Bicine-KOH, pH 8.3, 10 mM $MgCl_2$, 5 mM DTT, 1 mg ml^{-1} BSA, 1 mM ADP, 0.2 mM ATP, 0.2 mM P^1,P^5-di(adenosine-5′)pentaphosphate (Ap5A) in a final volume of 200 μl.

After 30 min at 30°C, 100 μl of the RP PK reaction was quenched with 100 μl SDS-PAGE sample buffer. The remaining 100 μl volume was rapidly desalted to remove interfering first-step PK assay metabolites. The second-step PP reaction was then initiated by addition of KPi and affinity purified AtRP (to final concentrations of 2.5 mM and approximately 0.2 μg ml^{-1}, respectively) and incubated for 30 min at 30°C prior to quenching the reaction with SDS-PAGE sample buffer as above. Aliquots (10 μl) of the quenched reactions were loaded onto 10% SDS-PAGE gels, electrophoresed and electroblotted onto a polyvinylidene fluoride (PVDF) membrane. The resulting blots were hybridized with anti-PPDK or anti-PPDK-ThrP polyclonal antibodies for immunoblot analysis as previously described (Chastain *et al.*, 2000, 2002, 2006). Specifically, the following primary antibodies used in this study were: (i) affinity-purified rabbit polyclonal antibodies raised against a synthetic phosphopeptide conjugate corresponding to the Thr-phosphorylation domain of maize C4 PPDK (Chastain *et al.*, 2000, 2002, 2006); (ii) affinity-purified rabbit polyclonal antibodies raised against maize recombinant C4 PPDK (Chastain *et al.*, 2006); and (iii) His_6-Tag monoclonal antibody (Novagen). Chemiluminescent detection of the respective antigen–antibody complexes on immunoblots was accomplished using the alkaline phosphatase substrate CDP-Star® (Applied Biosystems, http://www.appliedbiosystems.com/).

The spectrophotometric-based RP assays used for this study were essentially identical to the two-step RP PK and PP assay above except that unquenched aliquots (8 μl) of the incubating reactions were removed at 0 min and 20 min and assayed for PPDK activity using an enzyme-coupled spectrophotometric assay procedure as described previously (Chastain *et al.*, 2006).

Construction of RP1/2-GFP and CFP-RP2 fusion plasmids and microprojectile bombardment

Translational fusions between GFP and the protein-encoding *AtRP1/2* cDNAs were generated as follows. The deduced full-length ORFs of *AtRP1/2* were amplified from plasmid cDNA using Phusion® polymerase (New England Biolabs, http://www.neb.com/nebecomm/) with the following primer pairs: *AtRP1* F 5′-AATGGATCCATGGCTTTGCTCTCGGCGATGAAG-3′, *AtRP1* R 5′-AATGATATCATAGTAGCTTTTAGAGATGCGAGGC-3′; *AtRP2* F 5′- AATGGATCCATGAATTTACACGGCGTCTCGGGC-3′, *AtRP2* R 5′-AATGATATCATAGTAGCGTTTTGAGATACGCGG-3′. The primers introduced *Bam*HI and *Eco*RV sites at the 5′ and 3′ end of the amplification products, respectively, for ligation into a similarly cut pUC-based in-frame C-terminal fusion GFP vector that places expression under control of the CaMV-35S promoter (5′) and Nos terminator (3′) (Figure 4a,i). The constructs with only chloroplast-targeting peptide sequence predicted by ChloroP (Emanuelsson *et al.*,1999; Figure 4e,m) were constructed in the same manner except to amplify the respective truncated-length N-terminal ORF, and an alternative gene-specific reverse primer was used in the PCR reaction (*AtRP1* R 5′-AATGATATCTTCGATCGGACGCATCGGACCGGTC-3′; *AtRP2* R 5′-AATGATATCCAGTTCGAATTGTGGTCTAGAATTTC-3′). The N-terminal CFP full-length *AtRP2* translational fusion construct was generated in a similar manner except a *Sac*I site was introduced into the reverse *AtRP2* amplification primer (F 5′-ATGAATTTACACGGCGTCTCGGGC-3′, R 5′-AATGAGCTCTTAGTAGCGTTTTGAGATACGCGG-3′) for subsequent ligation into a vector containing the CaMV-35S promoter, CFP ORF and the Nos terminator (Figure 4q). Each GFP or CFP fusion construct was coated onto M-10 tungsten beads (Bio-Rad) and used in microprojectile bombardment of Arabidopsis (Columbia-0) or tobacco (*Nicotiana tabacum* cv. Petite Havana) leaves using the Bio-Rad PDS-1000/He particle delivery system (Hibberd *et al.*,1998). After overnight recovery, cells were visualized and digital images collected using a Leica SP confocal microscope (Leica Microsystems, http://www.leica-microsystems.com/). ADOBE PHOTOSHOP was used to overlay the chlorophyll and GFP or CFP channels.

Multiple alignments and molecular phylogenetic analysis

Bioinformatics and phylogenetic analyses of deduced aa sequences from isolated full-length *AtRP1/2* cDNAs, with representative Arabidopsis Ser/Thr PKs (AtCDPKs, AtPpcKs) and representative Arabidopsis Ser/Thr PP catalytic subunits was carried out with the CLUSTAL-X program (version 1.83; Thompson *et al.*, 1997). A distance matrix for the sequence alignment was calculated, and an unrooted tree was constructed using the PROTDIST (with the JTT model) and NEIGHBOR programs of the PHYLIP package (version 3.65), respectively (Felsenstein, 1996). Bootstrap analysis was performed with 1000 replications, and the tree was visualized using TREEVIEW 32 software. GenBank nucleotide accession numbers are listed in Table S1.

Acknowledgements

This work was supported by US National Science Foundation grants nos. IOS-0642190 to CJC and by MCB-0130057 (to RC) during CJC's sabbatical research in the Chollet laboratory. Additional funding was provided by UK Biotechnology and Biological Sciences Research Council grant P18931 and the Isaac Newton Trust to KP and JMH.

Supplementary Material

The following supplementary material is available for this article online:

Figure S1. The SDS–PAGE and immunoblot analysis of purified, recombinant AtRP1/2.

Table S1. Accession numbers of genes used for Figure 6 multiple alignment and phylogenetic tree analysis.

This material is available as part of the online article from http://www.blackwell-synergy.com

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

Bhaduri, A. and Sowdhamini, R. (2005) Genome-wide survey of prokaryotic O-protein phosphatases. *J. Mol. Biol.* **352**, 736–752.

Burnell, J.N. and Chastain, C.J. (2006) Cloning and expression of maize-leaf pyruvate, Pi dikinase regulatory protein gene. *Biochem. Biophys. Res. Commun.* **345**, 675–680.

Burnell, J.N. and Hatch, M.D. (1983) Dark/light regulation of pyruvate,Pi dikinase in C4 plants: evidence that the same protein catalyses activation and inactivation. *Biochem. Biophys. Res. Commun.* **111**, 288–293.

Burnell, J.N. and Hatch, M.D. (1985) Regulation of C4 photosynthesis: purification and properties of the protein catalyzing ADP-mediated inactivation and Pi-mediated activation of pyruvate, Pi dikinase. *Arch. Biochem. Biophys.* **237**, 490–503.

Chastain, C.J. and Chollet, R. (2003) Regulation of pyruvate,orthophosphate dikinase by ADP/Pi-dependent reversible phosphorylation in C3 and C4 plants. *Plant Physiol. Biochem.* **41**, 523–532.

Chastain, C.J., Lee, M.E., Moorman, M.A., Shameekumar, P. and Chollet, R. (1997) Site-directed mutagenesis of maize recombinant C4-pyruvate, orthophosphate dikinase at the phosphorylatable target threonine residue. *FEBS Lett.* **413**, 169–173.

Chastain, C.J., Botschner, M., Harrington, G.E., Thompson, B.J., Mills, S.E., Sarath, G. and Chollet, R. (2000) Further analysis of maize C_4-pyruvate,orthophosphate dikinase phosphorylation by its bifunctional regulatory protein using selective substitutions of the regulatory Thr-456 and catalytic His-458 residues. *Arch. Biochem. Biophys.* **375**, 165–170.

Chastain, C.J., Fries, J.P., Vogel, J.A. *et al.* (2002) Pyruvate,orthophosphate dikinase in leaves and chloroplasts of C_3 plants undergoes light/dark-Induced reversible phosphorylation. *Plant Physiol.* **128**, 1368–1378.

Chastain, C.J., Heck, J.W., Colquhoun, T.A., Voge, D.G. and Gu, X.Y. (2006) Posttranslational regulation of pyruvate, orthophosphate dikinase in developing rice (*Oryza sativa*) seeds. *Planta*, **224**, 924–934.

Dennis, D.T. and Blakeley, S.D. (2000) Carbohydrate metabolism. In *Biochemistry & Molecular Biology of Plants* (Buchanan, B.B., Gruissem, W. and Jones, R.L., eds). Rockville: American Society of Plant Physiologists, pp. 630–675.

Emanuelsson, O., Nielsen, H. and von Heijne, G. (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* **8**, 978–984.

Felsenstein, J. (1996) Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. *Meth. Enzymol.* **266**, 418–427.

Friso, G., Giacomelli, L., Ytterberg, A.J., Peltier, J-B., Rudella, A., Sun, Q. and van Wijk, K.J. (2004) In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: new proteins, new functions, and a plastid proteome database. *Plant Cell*, **16**, 478–499.

Furbank, R.T., Chitty, J.A., Jenkins, C.L.D., Taylor, W.C., Trevanion, S.J., Caemmerer, S.V. and Ashton, A.R. (1997) Genetic manipulation of key photosynthetic enzymes in the C4 plant *Flaveria bidentis*. *Aust. J. Plant Physiol.* **24**, 477–485.

Hanks, S.K. and Hunter, T. (1995) The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9**, 576–596.

Hardie, D.G. (1999) Plant protein serine/threonine kinases: classification and functions. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 97–131.

Hartwell, J., Gill, A., Nimmo, G.A., Wilkins, M.B., Jenkins, G.I. and Nimmo, H.G. (1999) Phosphoenolpyruvate carboxylase kinase is a novel protein kinase regulated at the level of expression. *Plant J.* **20**, 333–342.

Hibberd, J.M., Linley, P.J., Khan, M.S. and Gray, J.C. (1998) Transient expression of green fluorescent protein in various plastid types following microprojectile bombardment. *Plant J.* **16**, 627–632.

Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., De Castro, E., Langendijk-Genevaux, P.S., Pagni, M. and Sigrist, C.J.A. (2006) The PROSITE database. *Nucleic Acids Res.* **34**, D227–D230.

Lasanthi-Kudahettige, R., Magneschi, L., Loreti, E., Gonzali, S., Licausi, F., Novi, G., Beretta, O., Vitulli, F., Alpi, A. and Perata, P. (2007) Transcript profiling of the anoxic rice coleoptile. *Plant Physiol.* **144**, 218–231.

Leonard, C.J., Aravind, L. and Koonin, E.V. (1998) Novel families of putative protein kinases in bacteria and archaea: evolution of the 'eukaryotic' protein kinase superfamily. *Genome Res.* **8**, 1038–1047.

Majeran, W., Cai, Y. and van Wijk, K.J. (2005) Functional differentiation of bundle sheath and mesophyll maize chloroplasts determined by comparative proteomics. *Plant Cell*, **17**, 3111–3140.

Méchin, V., Thévenot, C., Le Guilloux, M., Prioul, J-L. and Damerval, C. (2007) Developmental analysis of maize endosperm proteome suggests a pivotal role for pyruvate orthophosphate dikinase. *Plant Physiol.* **143**, 1203–1219.

Parsley, K. and Hibberd, J.M. (2006) The *Arabidopsis* PPDK gene is transcribed from two promoters to produce differentially expressed transcripts responsible for cytosolic and plastidic proteins. *Plant Mol. Biol.* **62**, 339–349.

Peltier, J-B., Emanuelsson, O., Kalume, D.E. *et al.* (2002) Central functions of the lumenal and peripheral thylakoid proteome of *Arabidopsis* determined by experimentation and genome-wide prediction. *Plant Cell*, **14**, 211–236.

Plaxton, W.C. and Podesta, F.E. (2006) The functional organization and control of plant respiration. *Crit. Rev. Plant Sci.* **25**, 159–198.

Roeske, C.A. and Chollet, R. (1987) Chemical modification of the bifunctional regulatory protein of maize leaf pyruvate, orthophosphate dikinase: evidence for two distinct active sites. *J. Biol. Chem.* **262**, 12575–12582.

Scheeff, E.D. and Bourne, P.E. (2005) Structural evolution of the protein kinase-like superfamily. *PLoS Comput. Biol.* **5**, 359–381.

Small, I., Peeters, N., Legeai, F. and Lurin, C. (2004) Predotar: a tool for rapidly screening proteomes for *N*-terminal targeting sequences. *Proteomics*, **4**, 1581–1590.

Smith, C.M., Duff, S.M.G. and Chollet, R. (1994a) Partial purification and characterization of maize-leaf pyruvate, orthophosphate dikinase regulatory protein: a low-abundance, mesophyll-chloroplast stromal protein. *Arch. Biochem. Biophys.* **308**, 200–206.

Smith, C.M., Sarath, G. and Chollet, R. (1994b) A simple, single-tube radioisotopic assay for the phosphorylation/inactivation activity of the pyruvate,orthophosphate dikinase regulatory protein. *Photosynth. Res.* **40**, 295–301.

Stern, A., Privman, E., Rasis, M., Lavi, S. and Pupko, T. (2007) Evolution of the metazoan protein phosphatase 2C superfamily. *J. Mol. Evol.* **64**, 61–70.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882.

Whelan, S., de Bakker, P.I.W., Quevillon, E., Rodriguez, N. and Goldman, N. (2006) PANDIT: an evolution-centric database of protein and associated nucleotide domains with inferred trees. *Nucleic Acids Res.* **34**, D327–D331.

Xu, W., Zhou, Y. and Chollet, R. (2003) Identification and expression of a soybean nodule-enhanced PEP-carboxylase kinase gene (*NE-PpcK*) that shows striking up-/down-regulation *in vivo*. *Plant J.* **34**, 441–452.

Xu, W., Sato, S.J., Clemente, T.E. and Chollet, R. (2007) The PEP-carboxylase kinase gene family in *Glycine max* (*GmPpcK1-4*): an in-depth molecular analysis with nodulated, non-transgenic and transgenic plants. *Plant J.* **49**, 910–923.

Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. and Gruissem, W. (2004) GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol.* **136**, 2621–2632.

Accession numbers: *AtRP1*, At4g21210.1; *AtRP2*, At3g01200; *Glycine max* RP (*GmRP1*) (EU180222); *Sorghum bicolor* RP (*SbRP1*) (EU180223).

